

SEQUENCE LISTING

(1) GENERAL INFORMATION:

- (i) APPLICANTS: MINETTI, CONCEICAO;
MICHON, FRANCIS;
PULLEN, JEFFREY K.;
POLDVINO-BODNAR, MARYELLEN;
LIANG, SHU-MEI;
TAI, JOSEPH Y.
- (ii) TITLE OF INVENTION: MODIFIED IMMUNOGENIC
PNEUMOLYSIN COMPOSITIONS AS VACCINES
- (iii) NUMBER OF SEQUENCES: 18
- (iv) CORRESPONDENCE ADDRESS:
 - (A) ADDRESSEE: MORGAN & FINNEGAN, L.L.P.
 - (B) STREET: 345 PARK AVENUE
 - (C) CITY: NEW YORK
 - (D) STATE: NEW YORK
 - (E) COUNTRY: USA
 - (F) ZIP: 10154
- (v) COMPUTER READABLE FORM:
 - (A) MEDIUM TYPE: FLOPPY DISK
 - (B) COMPUTER: IBM PC COMPATIBLE
 - (C) OPERATING SYSTEM: PC-DOS/MS-DOS
 - (D) SOFTWARE: MICROSOFT WORD 97
- (vi) CURRENT APPLICATION DATA:
 - (A) APPLICATION NUMBER: 09/120,044
 - (B) FILING DATE: 1998-07-21
 - (C) CLASSIFICATION:
- (vii) PRIOR APPLICATION DATA:
 - (A) APPLICATION NUMBER: 60/053,306
 - (B) FILING DATE: 1997-07-21
- (vii) PRIOR APPLICATION DATA:
 - (A) APPLICATION NUMBER: 60/073,456
 - (B) FILING DATE: 1998-02-02
- (viii) ATTORNEY/AGENT INFORMATION:
 - (A) NAME: DARRYL H. STEENSMA
 - (B) REGISTRATION NUMBER: 43,155
 - (C) REFERENCE/DOCKET NUMBER: 1758-4036US2
- (ix) TELECOMMUNICATION INFORMATION:
 - (A) TELEPHONE: (212) 758-4800
 - (B) TELEFAX: (212) 751-6849
 - (C) TELEX: 421792

(2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1413
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(vi) ORIGINAL SOURCE:
 (A) ORGANISM: *S. pneumoniae*

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

atggcaaata aagcagtaaa tgactttata ctagctatga	40
attacgataa aaagaaactc ttgaccatc agggagaaag	80
tattgaaaat cgtttcatca aagagggtaa tcagctaccc	120
gatgagtttg ttgttatcga aagaaagaag cggagcttgt	160
cgacaaatac aagtgatatt tctgtaacag ctaccaacga	200
cagtcgcctc tctctggag cacttctcgt agtggatgag	240
accttgttag agaataatcc cactcttctt gcggtcgac	280
gtgctccgat gacttatagt attgatttgc ctggtttggc	320
aagtagcgat agctttctcc aagtggaaga tcccagcaat	360
tcaagtgttc gcggagcggg aaacgatttg ttggctaagt	400
ggcatcaaga ttatgggtcag gtcaataatg tcccagctag	440
aatgcagtat gaaaaaatca cggctcacag catggaacaa	480
ctcaagggtca agtttggttc tgactttgaa aagacaggga	520
attctcttga tattgatttt aactctgtcc attcaggcga	560
aaagcagatt cagattgtta attttaagca gatttattat	600
acagtcagcg tagacgctgt taaaaatcca ggagatgtgt	640
ttcaagatac tgtaacggta gaggatttaa aacagagagg	680
aatttctgca gagcgtcctt tggcttatat ttcgagtgtt	720
gcttatgggc gccaaagtcta tctcaagttg gaaaccacga	760
gtaagagtga tgaagtagag gctgcttttg aagctttgat	800
aaaaggagtc aaggtagctc ctcagacaga gtggaagcag	840
attttggaca atacagaagt gaaggcgggt attttagggg	880
gcgaccaag ttcgggtgcc cgagttgtaa caggcaaggt	920

```

ggatatggta gaggacttga ttcaagaagg cagtcgcttt      960
acagcagatc atccaggctt gccgatttcc tatacaactt      1000
cttttttacg tgacaatgta gttgcgacct ttcaaaatag      1040
tacagactat gttgagacta aggttacagc ttacagaaac      1080
ggagatttac tgctggatca tagtggtgcc tatgttgccc      1120
aatattatat tacttggaat gaattatcct atgatcatca      1160
aggtaaggaa gtcttgactc ctaaggcttg ggacagaaat      1200
gggcaggatt taacggctca ctttaccact agtattcctt      1240
taaaagggaa tgttcgtaat ctctctgtca aaattagaga      1280
gtgtaccggg cttgcttggg aatggtggcg tacggtttat      1320
gaaaaaaccg atttgccact agtgcgtaag cggacgattt      1360
ctatttgggg aacaactctc tatccgcagg tagaagataa      1400

ggtagaaaat gac  1413

```

(2) INFORMATION FOR SEQ ID NO:2:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 1413
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *S. pneumoniae*

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

```

atggcaaata aagcagtaaa tgactttata ctagctatga      40
attacgatan aaanaaaactc ttgacccatc agggagaaaag      80
tattgaaaat cgtttcanca aagagggtaa tcagctaccc      120
gntgagtttg ttgntancga aagaaagaag cggagcttgt      160
cgacaaatac aagtgatatt nctgtancag ctaccnacga      200
cagtcgcctc tatcctggag cacttctcgt agtggatgag      240
accttgtnag agaataatcc cactcttctt gcggtngatc      280
gtgctccgat gacttatagt antgntttgc ctggtttggc      320

```

aagtagcgat agctttctcc aagtggaaga ncccagcaat	360
tcaagtgttc gcggagcggn anacgatttg ttggctaagt	400
ggcatcaaga ttatggctcag gtcaataatg tcccagctag	440
aangcagtat gaaaaaatna cggctcacag catggaacaa	480
ctcaagggtca agtttggttc tgactttgaa aagncaggga	520
attctcttga tattgatttt aactctgtcc attcaggnga	560
aaagcngatt cagattgtta atnttaagca gatttattat	600
acagtcagcg tagacgtgt taaaaatcca ggagatgtgt	640
ttcaagatac tgtaacggta gaggatttaa aacagagagg	680
aatttctgca gagcgtcctt tgggtctatat ttcgagngtt	720
gcttatgggc gccaaagtcta tctcaagttg gaaaccacga	760
gtangagtgn tgaagtagag gctgcttttg aagctttgat	800
aaaaggagtc aaggtagctc ctcagacaga gtggaagcag	840
attttggaaca atacagaagt gaaggcgggtt attttagggg	880
gcgaccaag ttcgggtgcc cgagttgtaa caggcaaggt	920
ggatatggta gaggacttga ttcaagaagg cagtcgcttt	960
acagcagatc atccaggctt gccgatttcc tatacaactt	1000
cttttttacg tgacaatgta gttgcgacct ttcaaaanag	1040
tacagactat gttgagacta aggttacagc ttacagaaac	1080
ggagatttac tgctggatca tagtggtgcc tatgttgccc	1120
aatattatat tacttggnat gaattatcct atgatcatca	1160
aggtaaggaa gtcttgactc ctaaggcttg ggacagaaat	1200
gggcaggatt tnacggctca ctttaccact agtattcctt	1240
taaaagggaa tgttcgtaat ctctctgtca aaattagaga	1280
gtgtaccggg cttgcntggg aatggtggcg tacggtttat	1320
gaaaaaacgg atttgccact agtgcgtaag cggacgattt	1360
ctatttgggg aacaactctc tatccncagg tagangataa	1400
ggtagaaaat gac	1413

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 471
 (B) TYPE: amino acid
 (C) STRANDEDNESS: unknown
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: amino acid

(vi) ORIGINAL SOURCE:

- (A) ORGANISM:
- S. pneumoniae*

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

```

Met Ala Asn Lys Ala Val Asn Asp Phe Ile Leu Ala
1           5           10
Met Asn Tyr Asp Lys Lys Lys Leu Leu Thr His Gln
15           20
Gly Glu Ser Ile Glu Asn Arg Phe Ile Lys Glu Gly
25           30           35
Asn Gln Leu Pro Asp Glu Phe Val Val Ile Glu Arg
40           45
Lys Lys Arg Ser Leu Ser Thr Asn Thr Ser Asp Ile
50           55           60
Ser Val Thr Ala Thr Asn Asp Ser Arg Leu Tyr Pro
65           70
Gly Ala Leu Leu Val Val Asp Glu Thr Leu Leu Glu
75           80
Asn Asn Pro Thr Leu Leu Ala Val Asp Arg Ala Pro
85           90           95
Met Thr Tyr Ser Ile Asp Leu Pro Gly Leu Ala Ser
100          105
Ser Asp Ser Phe Leu Gln Val Glu Asp Pro Ser Asn
110          115          120
Ser Ser Val Arg Gly Ala Val Asn Asp Leu Leu Ala
125          130
Lys Trp His Gln Asp Tyr Gly Gln Val Asn Asn Val
135          140
Pro Ala Arg Met Gln Tyr Glu Lys Ile Thr Ala His
145          150          155
Ser Met Glu Gln Leu Lys Val Lys Phe Gly Ser Asp
160          165

Phe Glu Lys Thr Gly Asn Ser Leu Asp Ile Asp Phe
170          175          180
Asn Ser Val His Ser Gly Glu Lys Gln Ile Gln Ile
185          190
Val Asn Phe Lys Gln Ile Tyr Tyr Thr Val Ser Val
195          200
Asp Ala Val Lys Asn Pro Gly Asp Val Phe Gln Asp
205          210          215
Thr Val Thr Val Glu Asp Leu Lys Gln Arg Gly Ile
220          225
Ser Ala Glu Arg Pro Leu Val Tyr Ile Ser Ser Val
230          235          240
Ala Tyr Gly Arg Gln Val Tyr Leu Lys Leu Glu Thr
245          250
Thr Ser Lys Ser Asp Glu Val Glu Ala Ala Phe Glu

```

```

      255                      260
Ala Leu Ile Lys Gly Val Lys Val Ala Pro Gln Thr
265                      270                      275
Glu Trp Lys Gln Ile Leu Asp Asn Thr Glu Val Lys
      280                      285
Ala Val Ile Leu Gly Gly Asp Pro Ser Ser Gly Ala
290                      295                      300
Arg Val Val Thr Gly Lys Val Asp Met Val Glu Asp
      305                      310
Leu Ile Gln Glu Gly Ser Arg Phe Thr Ala Asp His
315                      320
Pro Gly Leu Pro Ile Ser Tyr Thr Thr Ser Phe Leu
325                      330                      335
Arg Asp Asn Val Val Ala Thr Phe Gln Asn Ser Thr
      340                      345
Asp Tyr Val Glu Thr Lys Val Thr Ala Tyr Arg Asn
350                      355                      360
Gly Asp Leu Leu Leu Asp His Ser Gly Ala Tyr Val
      365                      370
Ala Gln Tyr Tyr Ile Thr Trp Asn Glu Leu Ser Tyr
375                      380
Asp His Gln Gly Lys Glu Val Leu Thr Pro Lys Ala
385                      390                      395
Trp Asp Arg Asn Gly Gln Asp Leu Thr Ala His Phe
      400                      405
Thr Thr Ser Ile Pro Leu Lys Gly Asn Val Arg Asn
410                      415                      420
Leu Ser Val Lys Ile Arg Glu Cys Thr Gly Leu Ala
      425                      430
Trp Glu Trp Trp Arg Thr Val Tyr Glu Lys Thr Asp
435                      440

Leu Pro Leu Val Arg Lys Arg Thr Ile Ser Ile Trp
445                      450                      455
Gly Thr Thr Leu Tyr Pro Gln Val Glu Asp Lys Val
      460                      465
Glu Asn Asp
470

```

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 471
- (B) TYPE: amino acid
- (C) STRANDEDNESS: unknown
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: amino acid

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *S. pneumoniae*

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

```

Met Ala Asn Lys Ala Val Asn Asp Phe Ile Leu Ala
1                      5                      10
Met Asn Tyr Asp Xaa Xaa Lys Leu Leu Thr His Gln
      15                      20

```

```

Gly Glu Ser Ile Glu Asn Arg Phe Xaa Lys Glu Gly
 25          30          35
Asn Gln Leu Pro Xaa Glu Phe Val Xaa Xaa Glu Arg
      40          45
Lys Lys Arg Ser Leu Ser Thr Asn Thr Ser Asp Ile
 50          55          60
Xaa Val Xaa Ala Thr Xaa Asp Ser Arg Leu Tyr Pro
      65          70
Gly Ala Leu Leu Val Val Asp Glu Thr Xaa Leu Glu
 75          80
Asn Asn Pro Thr Leu Leu Ala Val Asp Arg Ala Pro
 85          90          95
Met Thr Tyr Ser Xaa Xaa Leu Pro Gly Leu Ala Ser
      100          105
Ser Asp Ser Phe Leu Gln Val Glu Asp Pro Ser Asn
 110          115          120
Ser Ser Val Arg Gly Ala Xaa Xaa Asp Leu Leu Ala
      125          130
Lys Trp His Gln Asp Tyr Gly Gln Val Asn Asn Val
 135          140

Pro Ala Arg Xaa Gln Tyr Glu Lys Xaa Thr Ala His
 145          150          155
Ser Met Glu Gln Leu Lys Val Lys Phe Gly Ser Asp
      160          165
Phe Glu Lys Xaa Gly Asn Ser Leu Asp Ile Asp Phe
 170          175          180
Asn Ser Val His Ser Gly Glu Lys Xaa Ile Gln Ile
      185          190
Val Asn Xaa Lys Gln Ile Tyr Tyr Thr Val Ser Val
 195          200
Asp Ala Val Lys Asn Pro Gly Asp Val Phe Gln Asp
 205          210          215
Thr Val Thr Val Glu Asp Leu Lys Gln Arg Gly Ile
      220          225
Ser Ala Glu Arg Pro Leu Val Tyr Ile Ser Xaa Val
 230          235          240
Ala Tyr Xaa Arg Gln Val Tyr Leu Lys Leu Glu Thr
      245          250
Thr Ser Xaa Ser Xaa Glu Val Glu Ala Ala Phe Glu
 255          260
Ala Leu Ile Lys Gly Val Lys Val Ala Pro Gln Thr
 265          270          275
Glu Trp Lys Gln Ile Leu Asp Asn Thr Xaa Val Lys
      280          285
Ala Val Ile Leu Gly Gly Asp Pro Ser Ser Gly Ala
 290          295          300
Arg Val Val Thr Gly Lys Val Asp Met Val Glu Asp
      305          310
Leu Ile Gln Glu Gly Ser Arg Phe Thr Ala Asp His
 315          320
Pro Gly Leu Pro Ile Ser Tyr Thr Thr Ser Phe Leu
 325          330          335
Arg Asp Asn Val Val Ala Thr Phe Gln Asn Ser Thr
      340          345
Asp Tyr Val Glu Thr Lys Val Thr Ala Tyr Arg Asn
 350          355          360
Gly Asp Leu Leu Leu Asp His Ser Gly Ala Tyr Val
      365          370

```

```

Ala Gln Tyr Tyr Ile Thr Trp Xaa Glu Leu Ser Tyr
      375                      380
Asp His Gln Gly Lys Glu Val Leu Thr Pro Lys Ala
385      390                      395
Trp Asp Arg Asn Gly Gln Asp Leu Thr Ala His Phe
      400                      405
Thr Thr Ser Ile Pro Leu Lys Gly Asn Val Arg Asn
      410                      415                      420

Leu Ser Val Lys Ile Arg Glu Cys Thr Gly Leu Ala
      425                      430
Trp Glu Trp Trp Arg Thr Val Tyr Glu Lys Thr Asp
      435                      440
Leu Xaa Leu Val Arg Lys Arg Thr Ile Ser Ile Trp
445      450                      455
Gly Thr Thr Leu Tyr Pro Gln Val Glu Asp Lys Val
      460                      465
Glu Asn Asp
      470

```

(2) INFORMATION FOR SEQ ID NO:5:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 35
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(vi) ORIGINAL SOURCE:
 (A) ORGANISM: *S. pneumoniae*

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

aaccttgatt gatctagata aggtatttat gttgg 35

(2) INFORMATION FOR SEQ ID NO:6:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 33
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(vi) ORIGINAL SOURCE:
 (A) ORGANISM: *S. pneumoniae*

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

tctttttgtc tctagaattc tcctctccta gtc 33

(2) INFORMATION FOR SEQ ID NO:7:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 37
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

- (vi) ORIGINAL SOURCE:
 - (A) ORGANISM: *S. pneumoniae*

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

tattaggagg agcatatggc aaataaagca gtaaattg 37

(2) INFORMATION FOR SEQ ID NO:8:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 37
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

- (vi) ORIGINAL SOURCE:
 - (A) ORGANISM: *S. pneumoniae*

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

ggcctctttt tgtctcgagc attctcctct cctagtc 37

(2) INFORMATION FOR SEQ ID NO:9:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 21
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: nucleic acid

- (vi) ORIGINAL SOURCE:
 - (A) ORGANISM: *S. pneumoniae*

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

attacgcgac tcactatagg g 21

(2) INFORMATION FOR SEQ ID NO:10:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 21
 - (B) TYPE: nucleic acid

(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: *S. pneumoniae*

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

attacgaaca ttcccttttag g 21

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 37
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: *S. pneumoniae*

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

ggtcagggtca ataatgtccc agctagaaag cagtatg 37

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 38
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: *S. pneumoniae*

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

gctgtgagcc gtgatttttt catactgctt tctagctg 38

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 38
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: *S. pneumoniae*

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

gcagattcag attgttaatg ttaagcagat ttattata 38

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 38
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: *S. pneumoniae*

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

atctgcttaa cattaacaat ctgaatctgc ttttcgcc 38

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 39
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: *S. pneumoniae*

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

cagattgtta atattaagca gatttattat acagtcagc 39

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 39
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: *S. pneumoniae*

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

aatctgctta atattaacaa tctgaatctg cttttcgcc 39

(2) INFORMATION FOR SEQ ID NO:17:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 39
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *S. pneumoniae*

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

acaagtgata ttctgtaac agctaccaac gacagtcgc 39

(2) INFORMATION FOR SEQ ID NO:18:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 39
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *S. pneumoniae*

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

agctgttaca ggaatatcac ttgtatttgc cgacaagct 39